

一种针对代谢网络多平衡态性质的 强连通分解方法^{*}

毕文健 郭金 赵延龙 张纪峰

(中国科学院数学与系统科学研究院系统科学研究所系统控制重点实验室, 北京 100190)

摘要 针对反应速率满足一定条件的代谢网络, 提出了一种强连通分解方法对网络进行分解, 通过研究分解后的子网络来分析整体网络的多平衡态性质. 基于代谢网络的拓扑结构, 构造了其对应的代谢反应图和相互作用图, 引入了紧缩运算的定义, 构造了强连通分解算法; 给出了该算法的计算复杂度, 证明了分解的唯一性以及分解后子网络的强连通性, 阐明了子网络与整体网络在多平衡态性质意义下的关系, 举例说明了强连通算法和所得主要结果的有效性.

关键词 代谢网络, 强连通分解算法, 相互作用图, 多平衡态性质, 紧缩运算.

MR(2000) 主题分类号 92C42

1 引言

对代谢网络的研究是系统生物学的一个重要分支^[1], 其中代谢网络的多平衡态性质问题是一个基本且意义重大的科学问题^[2-6]. 比如现代分子生物学认为: 对基因调控网络的多平衡态的研究可以解释细胞分化现象^[7]; 又如在农业中, 植物叶肉细胞将二氧化碳转化为能量物质的光合碳代谢网络的多平衡态性质与粮食作物的产量密切相关^[8]. 因此, 代谢网络的多平衡态问题逐渐受到人们的广泛关注.

通过实验方法研究多平衡态问题不仅花费巨大, 而且由于实验条件所限常常无法实现^[5], 通常需要借用数学模型对其进行理论分析^[9-11]. 在数学模型的分析中, 往往会需要大量参数信息. 这些参数不仅不易获取, 而且有时还会随着外界环境和自身条件的变化而改变^[12-13]. 数学模型的参数不确定性会导致理论结果的局限性. 相对于参数信息, 代谢网络的拓扑结构更易获得也相对稳定^[5]. 因此, 基于拓扑结构对代谢网络进行研究不仅能够克服参数信息不足带来的困难, 而且具有更大的普适性.

^{*} 国家自然科学基金 (61134013, 11171333), 中国科学院青年促进会基金 (4106960) 资助课题.
收稿日期: 2011-01-30, 收到修改稿日期: 2012-05-08.

代谢网络的拓扑结构与功能之间的关系早已受到学者们的关注, 其中针对多平衡态性质的工作也有很多^[3,14-17]. 文 [3] 提出了“模块化”的研究思路, 证明了无抑制作用的单底物单产物 (SSN) 代谢模块在含有输出节点时不存在多个平衡态; 文 [17] 进一步分析了多底物多产物无抑制作用 (MMN) 的代谢模块, 得到了不存在多个平衡态的充分条件, 并给出了一类不存在多平衡态的特殊代谢模块: \mathcal{A} -MMN 模块. 这些工作在做理论分析时, 大多是利用单射性作为不存在多个平衡态的充分条件. 然而, 对于不满足单射条件的代谢网络, 无法分析其多平衡态性质.

除了利用单射性之外, 对代谢网络的分解也是一个有效的理论分析方法. 考虑到代谢网络节点众多且结构复杂, 许多工作将其分解成为若干相对简单的子网络^[12-13,18-21]. 但这些工作大多是通过某些指标得到一些具有功能独立性的子网络, 而不是通过分析子网络的性质来达到研究整体网络的目的.

本文针对代谢网络的多平衡态性质提出了一种强连通分解算法, 将一个整体网络分解为多个强连通的子网络. 证明了该算法是多项式算法且满足分解唯一性. 然后证明了如果代谢反应速率满足一定条件, 子网络与整体网络在多平衡态性质意义下存在如下关系

i) 如果分解后的每一个子网络都不具有多平衡态能力, 则整体网络也不具有多平衡态能力;

ii) 如果存在一个子网络可以产生多平衡态且整体网络中的其它子网络可以传递这种多平衡态性质, 则整体网络具有多平衡态能力;

iii) 如果存在一个子网络不能产生平衡态, 则整体网络不能产生平衡态;

iv) 如果每一个子网络对于多平衡态有传递能力, 则整体网络也具有这种传递能力.

本文结构如下: 第 2 节描述代谢网络的有向图绘制及多平衡态性质问题; 第 3 节介绍本文需要的一些预备知识; 第 4 节给出强连通分解算法, 证明算法的合理性和可行性, 分析分解后的子网络与整体网络在平衡态性质意义下的关系; 第 5 节用几个例子来说明本文得到的一些结果; 第 6 节对本文的工作进行总结与展望.

2 问题描述

对给定的代谢网络, 用 \mathcal{M} 表示所有代谢物组成的集合, \mathcal{R} 表示所有代谢反应组成的集合, 并把这个代谢网络记为 $(\mathcal{M}, \mathcal{R})$ ^[17].

定义 2.1^[17] $(\mathcal{M}, \mathcal{R})$ 的代谢反应图是把代谢物看成节点, 连接代谢物之间的反应看成边, 方向由反应物指向生成物. 对于一个节点, 如果所有与之相连的边都由该节点指向其它节点, 则称之为输入节点, 如果所有与之相连的边都指向该节点, 则称之为输出节点, 除输入节点和输出节点之外的节点称为状态节点.

设 $(\mathcal{M}, \mathcal{R})$ 有 l 个输入节点, m 个状态节点和 k 个代谢反应. 以加中括号的状态节点名称表示其浓度, $s = ([S_1], [S_2], \dots, [S_m])^T$, 可以得到 $(\mathcal{M}, \mathcal{R})$ 的动态模型^[17,23] 为

$$\frac{ds}{dt} = r(s; \mu, \nu) = Mv(s; \mu, \nu), \quad (1)$$

其中输入节点的浓度 $\mu \in \mathbb{R}_+^l$ 看做模型参数, 称为输入参量. 用 $\nu \in V$ 表示输入节点以外的模型参数, 称为系统参量, V 表示 ν 所有可能取值组成的集合. M 表示代谢网络的化学

计量矩阵, v 表示 \mathcal{R} 中的各个代谢反应的速率按照一定顺序组成的 k 维向量. 化学反应速率 v 与模型参数和状态节点浓度相关, 而化学计量矩阵 $M = (M_{ij})_{1 \leq i \leq m, 1 \leq j \leq k}$ 只体现代谢网络的拓扑结构.

注 1 若一个代谢反应是可逆的, 则将其视为两个反应: 正反应和逆反应. 例如将可逆反应 $A + B \rightleftharpoons C + D$ 分成正反应 $A + B \rightarrow C + D$ 与逆反应 $C + D \rightarrow A + B$.

假设 2.1 对于 (M, \mathcal{R}) 中的反应 R_j ($1 \leq j \leq k$), 它的速率 v_j 满足下面假设

$$\begin{cases} \frac{\partial v_j}{\partial [S_i]} > 0, & M_{ij} < 0, \\ \frac{\partial v_j}{\partial [S_i]} = 0, & M_{ij} \geq 0, \end{cases}$$

其中 $[S_i]$ ($1 \leq i \leq m$) 为第 i 个节点的浓度.

注 2 假设 2.1 中对于反应速率的描述与质量作用定律 (Law of mass action)^[24], M-M (Michaelis-Menten) 方程^[25], 希尔 (Hill) 机制^[26] 是一致的.

定义 2.2 (M, \mathcal{R}) 的相互作用图 $G(V, E)$ 是有向图, 其节点集 V 对应于状态节点集 $S = \{S_1, S_2, \dots, S_m\}$, 有向弧集 E 由以下方式构造

$$E = \{\text{弧 } S_i S_j : i \neq j, 1 \leq i \leq m, 1 \leq j \leq m, \text{ 且 } \exists 1 \leq t \leq k, \text{ 使得 } M_{it} < 0, M_{jt} \neq 0\}.$$

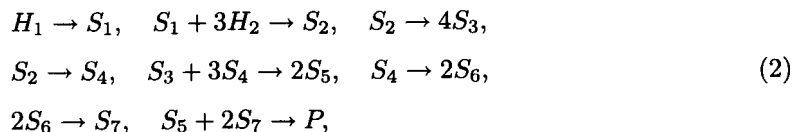
引理 2.1 如果代谢网络 (M, \mathcal{R}) 满足假设 2.1, 且其相互作用图 G 不存在弧 $S_i S_j$, $1 \leq i \leq m, 1 \leq j \leq m$, 则 $\frac{\partial r_j}{\partial [S_i]} = 0$, 其中 r_j 代表 r 的第 j 个分量.

证 对 $1 \leq i \leq m, 1 \leq j \leq m$, 由式 (1) 知

$$r_j = \sum_{t=1}^k M_{jt} v_t \Rightarrow \frac{\partial r_j}{\partial [S_i]} = \sum_{t=1}^k M_{jt} \frac{\partial v_t}{\partial [S_i]}.$$

因为 G 不存在弧 $S_i S_j$, 由定义 2.2, $\forall 1 \leq t \leq k$, 若 $M_{jt} \neq 0$, 则 $M_{it} \geq 0$. 而由假设 2.1, $M_{it} \geq 0 \Rightarrow \frac{\partial v_t}{\partial [S_i]} = 0$, 故 $\forall 1 \leq t \leq k, M_{jt} \frac{\partial v_t}{\partial [S_i]} = 0$, 即 $\frac{\partial r_j}{\partial [S_i]} = 0$.

例 2.1 下面举例解释前面的概念: 考虑代谢网络



其代谢反应图如图 1 所示, 相互作用图如图 2 所示. H_1 和 H_2 是输入节点, P 是输出节点, 其它节点是状态节点.

下面根据代谢网络 (M, \mathcal{R}) 具有的不同平衡态性质对其进行分类.

定义 2.3 对任意的 $\mu \in \mathbb{R}_+^l$ 和 $\nu \in V$, 如果代数方程 $r(s; \mu, \nu) = 0$ 关于 s 在 \mathbb{R}_+^m 中都至多有一个解, 那么称 (M, \mathcal{R}) 没有能力产生多平衡态, 如果 $r(s; \mu, \nu) = 0$ 都没有解, 那么称 (M, \mathcal{R}) 没有能力产生平衡态.

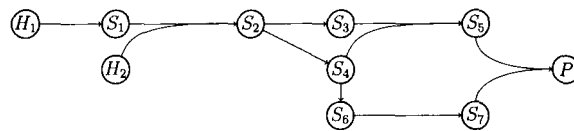


图 1 代谢网络的代谢反应图

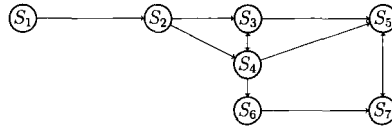


图 2 代谢网络的相互作用图

定义 2.4 对任意的 $\mu \in \mathbb{R}_+^l$, 如果存在 $\nu \in V$, 使得 $r(s; \mu, \nu) = 0$ 关于 s 在 \mathbb{R}_+^m 中存在至少一个解, 那么称 (M, \mathcal{R}) 对于任意输入能维持平衡态, 如果存在 $\nu \in V$, 使得 $r(s; \mu, \nu) = 0$ 在 \mathbb{R}_+^m 中至少存在两个解, 那么称 (M, \mathcal{R}) 对于任意输入能产生多平衡态.

本文的目的是代谢网络在多平衡态性质意义下的分解, 以期分解后的子网络可以反映甚至决定整体网络的性质. 为此, 需要设计子网络的分解算法, 论证算法的合理性和可行性, 并揭示子网络与整体网络的关系.

3 预备知识

为了叙述的方便与严谨, 这里介绍一些图论中的概念和结论^[22].

定义 3.1 若有向图 $G(V, E)$ 中存在一个有向边由节点 S_i 指向节点 S_j , 则认为 G 中存在弧 $S_i S_j$. 有向图中的一个途径是指一个由节点和弧交错排列的序列 $W_{i,j} = S_i a_i S_{i+1} \cdots S_{j-1} a_{j-1} S_j$, 使得 $\forall i \leq t \leq j-1$, 弧 a_t 即弧 $S_t S_{t+1}$, 称 $W_{i,j}$ 是一条从 S_i 到 S_j 的 (S_i, S_j) 途径. 如果有向图中存在一条 (S_i, S_j) 途径, 则称节点 S_j 从 S_i 是可达的.

定义 3.2 一个有向图 G 的任意一对节点 S_i 与 S_j 相互可达, 则称 G 是强连通的.

定义 3.3 有向图 G 中 j 个不同的节点 S_1, S_2, \dots, S_j 构成了一个圈, 如果 G 中含有弧 $S_1 S_2, S_2 S_3, \dots, S_{j-1} S_j, S_j S_1$. 此时称节点 S_1, S_2, \dots, S_j 是圈中的节点, 其它节点为该圈外的节点.

定义 3.4 节点的入度为指向这个节点的弧的数量, 节点的出度为由该节点指出的弧的数量.

定义 3.5 对给定的无圈有向图 G 和它的一个节点序 S_1, S_2, \dots, S_n , 若对 G 中的每一条弧 $S_i S_j$, 总有 $i < j$, 则称这个序是无圈序.

命题 3.1^[22] 一个有向图如果节点的入度都不为零, 则存在圈.

命题 3.2^[22] 任何无圈有向图 G 都含有无圈序. 如果 G 中含有 m 个节点, r 条弧, 则构造无圈序的过程能够在 $O(m+r)$ 步内完成.

下面定义一种具有特殊结构的代谢网络.

定义 3.6 若一个代谢网络形成的相互作用图是一个强连通图, 则称该代谢网络为强连

通网络.

4 代谢网络的强连通分解

本节将给出代谢网络的一种直观的强连通分解算法, 并阐明分解后的子网络与整体网络在多平衡态性质下的关系.

4.1 强连通分解算法

介绍强连通分解算法之前, 先引入紧缩运算的定义, 并在图 3 中用一个简单例子解释.

定义 4.1 有向图 G 中含有圈 C , 对 C 的紧缩运算由如下 3 步组成.

- i) 删除节点和弧: 删除圈 C 中的节点以及与它们相连的弧;
- ii) 添加节点: 添加一个新节点 V' 作为圈 C 的整体节点;
- iii) 添加弧: 如果 G 中存在由圈外节点 $V_{i,1}$ 指向圈中节点 $V_{j,1}$ 的弧, 则添加弧由 $V_{i,1}$ 指向 V' ; 如果原有向图中存在由圈中节点 $V_{j,2}$ 指向圈外节点 $V_{i,2}$ 的弧, 则添加弧由 V' 指向 $V_{i,2}$.

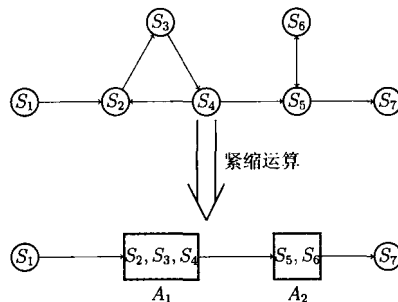


图 3 有向图的紧缩运算

(紧缩运算把 S_2, S_3, S_4 紧缩成节点 A_1 , 把 S_5, S_6 紧缩成节点 A_2)

强连通分解算法

第 1 步 绘制相互作用图

由定义 2.2 绘制代谢网络对应的相互作用图 G .

第 2 步 检测相互作用图中是否有圈

删除图 G 中入度为零的节点生成一个新有向图. 重新计算剩余各节点入度, 再删除入度为零的节点, 以此类推. 经过有限步删除后,

- i) 如果全部节点被删除, 则图 G 不存在圈, 进入第 4 步.
- ii) 如果剩下若干节点无法继续删除, 即它们的入度都不为零, 由命题 3.1, 图 G 存在圈, 进入第 3 步.

第 3 步 对有圈有向图进行简化

对有向图中的圈进行紧缩运算, 然后回到第 2 步继续检测.

第 4 步 对无圈有向图进行排序

经过有限步紧缩, 图 G 中的圈被全部删除, 得到一个无圈有向图 G' . 图 G' 中的节点 (称为阶段) 由原有向图 G 中的有限个节点 (称为阶段包含的节点) 经过紧缩运算生成. 根据命题 3.2, 无圈有向图 G' 存在无圈序. 假设排序后有 n 个阶段 A_1, A_2, \dots, A_n , 其中 A_i 包含状态节点 $S_{i,1}, S_{i,2}, \dots, S_{i,m_i}$, 则找到一种对于状态节点的分解方式

$$A = A_1 \cup A_2 \cup \dots \cup A_n.$$

第 5 步 生成代谢子网络

对于阶段 A_i , 将 $S_{i,1}, S_{i,2}, \dots, S_{i,m_i}$ 参与的全部代谢反应集合记为 \mathcal{R}_i , 把 \mathcal{R}_i 中涉及的全部代谢物质集记为 \mathcal{M}_i , 就得到代谢子网络 $(\mathcal{M}_i, \mathcal{R}_i)$. 整体网络 $(\mathcal{M}, \mathcal{R})$ 被分解成为多个子网络.

下面说明强连通分解算法的合理性和可行性.

定理 4.1 如果代谢网络 $(\mathcal{M}, \mathcal{R})$ 满足假设 2.1, 且相互作用图中含有 m 个节点, r 个弧, 那么用强连通分解算法分解此网络的时间复杂度为 $O(m^2 + r)$.

证 在分解算法中, 由于有向图中含有 m 个节点, r 个弧, 第 2 步最多进行 m 次节点入度的计算与删除, 故复杂度为 $O(m)$. 而有向图中每次进行紧缩运算后得到的新图至少比原图少一个节点, 故第 2 步与第 3 步之间最多经历 m 次循环. 由命题 3.2, 排序算法复杂度为 $O(m + r)$. 故整个分解算法的复杂度为 $m \cdot O(m) + O(m + r) = O(m^2 + r)$, 即整个分解算法是一个多项式算法.

定理 4.2 在定理 4.1 的条件下, 对 $(\mathcal{M}, \mathcal{R})$ 进行强连通分解后, 得到的子网络都是强连通的.

证 由于圈中节点相互可达, 由定义 4.1 知有限步紧缩运算生成的整体节点 $A_i, 1 \leq i \leq n$ 中包含的状态节点 $S_{i,1}, S_{i,2}, \dots, S_{i,m_i}$ 也相互可达, 这等价于相互作用图的强连通性. 由定义 3.6 知子网络是强连通的.

定理 4.3 在定理 4.1 的条件下, $(\mathcal{M}, \mathcal{R})$ 的强连通分解具有分解唯一性.

证 定理 4.2 证明强连通分解生成的每一个阶段对应的有向图是强连通图. 而阶段内包含的节点不可能与阶段外的节点相互可达, 否则一定存在圈, 可以再进行一次紧缩, 故是一个能扩充的最大的强连通图. 由文 [22] 可知这种扩充生成的强连通图是唯一的, 进而强连通分解也具有分解唯一性.

注 3 有向图的强连通分解算法已有很多, 比如深度优先搜索的 Tarjan 算法 [27] 等. 已有算法大多以数据结构知识为基础, 常常涉及搜索树、栈、十字链表等概念. 相比这些搜索算法, 本文提出的分解算法只基于最基本的图论知识, 易于理解, 便于操作, 有利于生物学者掌握与使用.

4.2 强连通子网络与整体网络的关系

本小节将给出分解后的子网络与整体网络在多平衡态性质意义下的关系 (具体见表 1).

表 1 强连通子网络与整体网络对应关系

定理	强连通子网络性质	整体网络性质	实例
定理 4.4	每个子网络都没有能力产生多平衡态	没有能力产生多平衡态	例 5.2
定理 4.5	存在至少一个子网络没有能力产生平衡态	没有能力产生平衡态	例 5.1
定理 4.6	存在至少一个子网络对任意输入能产生多平衡态, 其他子网络对任意输入能维持平衡态	对于任意输入能产生多平衡态	例 5.3
定理 4.7	每个子网络都对任意输入能维持多平衡态	对于任意输入能维持多平衡态	例 5.2

假设 (M, \mathcal{R}) 经过强连通分解被分成 n 个子网络, 状态节点集 $A = \{S_{1,1}, S_{1,2}, \dots, S_{1,m_1}, S_{2,1}, S_{2,2}, \dots, S_{2,m_2}, \dots, S_{n,1}, S_{n,2}, \dots, S_{n,m_n}\}$ 相应地被分为 n 个阶段: $A = A_1 \cup A_2 \cup \dots \cup A_n$, 其中 $A_i = \{S_{i,1}, S_{i,2}, \dots, S_{i,m_i}\}, \forall i = 1, 2, \dots, n$. 系统参量 ν 也分为 n 部分: 若含有 ν_m 的 $v_j(s; \mu, \nu)$ 对应的反应 R_j 属于 (M_i, \mathcal{R}_i) , 则称系统参量 ν_m 属于子网络 (M_i, \mathcal{R}_i) . 用向量 ν_i 表示属于 (M_i, \mathcal{R}_i) 的系统参量.

由强连通分解算法可以得到如下引理.

引理 4.1 代谢网络 (M, \mathcal{R}) 经过强连通分解, 状态节点集 A 被分为 n 个阶段: $A = A_1 \cup A_2 \cup \dots \cup A_n$. $\forall i > j$, 任取 α, β , 使得 $S_\alpha \in A_j, S_\beta \in A_i$, 则 $\frac{\partial r_\alpha}{\partial [S_\beta]} = 0$.

证 由 A_i 与 A_j 之间的序关系可知, A_i 中节点没有到 A_j 中节点的途径. 进而由引理 2.1 可知结论成立.

现证明本文最核心的 4 个定理, 这些定理说明可以通过分析子网络得到整体网络的平衡态性质. 而由于子网络结构相对简单, 故对子网络的分析也会相对容易.

证明中假设将整体网络 (M, \mathcal{R}) 强连通分解为 n 个子网络 $(M_i, \mathcal{R}_i), 1 \leq i \leq n$.

定理 4.4 在定理 4.1 的条件下, 如果 $\forall i \leq n, (M_i, \mathcal{R}_i)$ 都没有能力产生多平衡态, 则整体网络 (M, \mathcal{R}) 也没有能力产生多平衡态.

证 对于阶段 A_i 对应的子网络 (M_i, \mathcal{R}_i) , 其系统参量为 ν_i . 用 s_i 表示 (M_i, \mathcal{R}_i) 中状态节点浓度, 即 $s_i = [[S_{i,1}], [S_{i,2}], \dots, [S_{i,m_i}]]^T$. 由分解的序关系可知, 它的输入参量由两部分组成: i) 整体系统的输入参量 μ_0 ; ii) 阶段 A_1 到 A_{i-1} 的状态变量 $s_j, 1 \leq j < i$. 记为 $\mu_i = [\mu_0^T, s_1^T, \dots, s_{i-1}^T]^T$, 则由引理 4.1, 原有的状态方程 (1) 可以改为下方方程 (3) 形式, 其中每一个等式都可以看成子网络的状态方程

$$\begin{cases} \frac{ds_1}{dt} = r_1(s_1; \mu_1, \nu_1) \\ \frac{ds_2}{dt} = r_2(s_2; \mu_2, \nu_2) \\ \vdots \\ \frac{ds_i}{dt} = r_i(s_i; \mu_i, \nu_i) \\ \vdots \\ \frac{ds_n}{dt} = r_n(s_n; \mu_n, \nu_n) \end{cases} \quad (3)$$

反证法: 若整体网络在输入为 μ_0 , 参数为 ν 时存在两个平衡态, 即存在 $s \neq s'$, 使得 $r(s; \mu_0, \nu) = r(s'; \mu_0, \nu) = 0$. 不失一般性, 假设存在 $p, s_1 = s'_1, s_2 = s'_2, \dots, s_{p-1} = s'_{p-1}$,

$s_p \neq s'_p$. 针对 A_p 这一子网络, 保持 μ_p 与 ν_p 不变, 由 $r_p(s_p; \mu_p, \nu_p) = r_p(s'_p; \mu_p, \nu_p) = 0$, 可知子网络 (M_p, \mathcal{R}_p) 在参数 μ_p 与 ν_p 下存在两个平衡态 $s_p \neq s'_p$, 与条件矛盾. 所以如果每一个阶段都没有能力产生多个平衡态, 则整个系统也一定没有能力产生多个平衡态.

注 4 由定理 4.4 知整个系统的多平衡态由至少一个阶段提供.

定理 4.5 在定理 4.1 的条件, 若 $\exists i \leq n$ 子网络 (M_i, \mathcal{R}_i) 没有能力产生平衡态, 则整体网络 (M, \mathcal{R}) 没有能力产生平衡态.

证 反证法: 假设 (M, \mathcal{R}) 存在平衡态, 即在参数 ν 和 μ_0 下存在 s , 使得 $r(s; \mu_0, \nu) = 0$. 那么, 对于所有 i , $r_i(s_i; \mu_i, \nu_i) = 0$, 这等价于子网络 (M_i, \mathcal{R}_i) 在参数 ν_i 和 μ_i 下存在平衡态 s_i . 由 i 的任意性可知, 所有子网络都存在平衡态. 所以如果存在子网络没有能力产生平衡态, 那么整个网络也没有能力产生平衡态.

定理 4.6 在定理 4.1 的条件下, 强连通分解后存在至少一个子网络对于任意输入能产生多平衡态, 而其它子网络也对于任意输入能维持平衡态, 那么整个网络对于任意输入能产生多平衡态.

证 不失一般性, 假设子网络 (M_i, \mathcal{R}_i) 对于任意输入能产生多平衡态, 其它子网络对于任意输入能维持平衡态. 因为子网络 (M_1, \mathcal{R}_1) 对于任意输入能维持平衡态, 所以任意给定 μ_1 , 存在系统参量 ν_1 , 方程 $r_1(s_1; \mu_1, \nu_1) = 0$ 有解 s_1 . 因为子网络 (M_2, \mathcal{R}_2) 对于任意输入能维持平衡态, 所以对于输入参量 $\mu_2 = [\mu_1^T, s_1^T]^T$, 存在系统参数 ν_2 , 使方程 $r_2(s_2; \mu_2, \nu_2) = 0$ 有解 s_2 . 依次类推, 对于子网络 (M_i, \mathcal{R}_i) , 因为其对于任意输入能产生多平衡态, 所以对于输入参量 $\mu_i = [\mu_{i-1}^T, s_{i-1}^T]^T$, 存在系统参量 ν_i , 使方程 $r_i(s_i; \mu_i, \nu_i) = 0$ 有两个解 s_i 和 s'_i . 而由于子网络 $(M_{i+1}, \mathcal{R}_{i+1})$ 对于任意输入能维持平衡, 所以对于两组不同的输入参量 $\mu_{i+1} = [\mu_i^T, s_i^T]^T$ 和 $\mu'_{i+1} = [\mu_i^T, s'_i{}^T]^T$, 分别取不同的系统参量 ν_{i+1} 和 ν'_{i+1} , 都可以使方程 $r_{i+1}(s_{i+1}; \mu_{i+1}, \nu_{i+1}) = 0$ 有解. 以此类推可知, 整个网络也是对于任意输入能产生多平衡态.

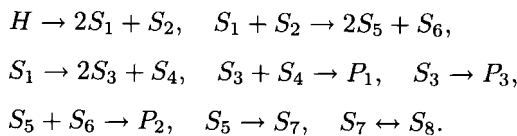
定理 4.7 在定理 4.1 的条件下, 若 (M, \mathcal{R}) 强连通分解后的每一个子网络都对于任意输入能维持多平衡态, 则整个网络也对于任意输入能维持多平衡态.

证 由定理 4.6 的前半部分可知结论成立.

5 例子

假定本节例子中反应速率服从 M - M 方程形式^[25].

例 5.1 考虑代谢网络



由强连通分解可以得到如下 4 个子网络

$$(\mathcal{M}_{1,1}, \mathcal{R}_{1,1}): H \rightarrow 2S_1 + S_2, S_1 + S_2 \rightarrow P_1, S_1 \rightarrow P_2;$$

$$(\mathcal{M}_{1,2}, \mathcal{R}_{1,2}): H \rightarrow 2S_3 + S_4, S_3 + S_4 \rightarrow P_1, S_3 \rightarrow P_2;$$

$$(\mathcal{M}_{1,3}, \mathcal{R}_{1,3}): H \rightarrow 2S_5 + S_6, S_5 + S_6 \rightarrow P_1, S_5 \rightarrow P_2;$$

$$(\mathcal{M}_{1,4}, \mathcal{R}_{1,4}): H \rightarrow S_7, S_7 \leftrightarrow S_8.$$

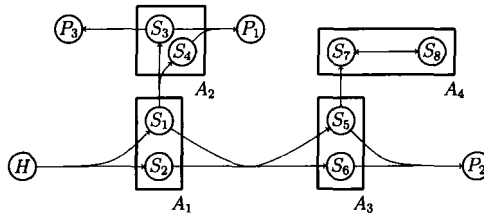


图 4 例 5.1 的强连通分解图例

对子网络 $(\mathcal{M}_{1,4}, \mathcal{R}_{1,4})$ 进行分析, 其状态方程为

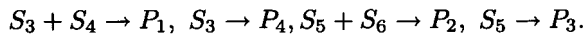
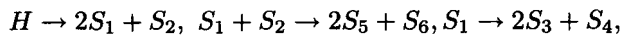
$$\begin{cases} \frac{d[S_7]}{dt} = v_1 - v_2 + v_3, \\ \frac{d[S_8]}{dt} = v_2 - v_3, \end{cases}$$

其中 $v_1 = u$, $v_2 = \frac{V_1[S_7]}{k_1 + [S_7]}$, $v_3 = \frac{V_2[S_8]}{k_2 + [S_8]}$.

由于 $v_1 = u > 0$, 所以状态方程右端不为零, 即该子网络没有能力产生平衡态. 由定理 4.5 可知, 整个网络也没有能力产生平衡态.

注 5 例 5.1 体现出强连通分解的优势. 通过强连通分解, 把一个含有 8 个状态节点的代谢网络转为 4 个子网络, 由只含有 2 个状态节点的子网络 $(\mathcal{M}_{1,4}, \mathcal{R}_{1,4})$ 的结构特性就得到了整体网络的多平衡态性质.

例 5.2 考虑代谢网络



由强连通分解可以得到如下 3 个子网络

$$(\mathcal{M}_{2,1}, \mathcal{R}_{2,1}): H \rightarrow 2S_1 + S_2, S_1 + S_2 \rightarrow P_1, S_1 \rightarrow P_2;$$

$$(\mathcal{M}_{2,2}, \mathcal{R}_{2,2}): H \rightarrow 2S_3 + S_4, S_3 + S_4 \rightarrow P_1, S_3 \rightarrow P_2;$$

$$(\mathcal{M}_{2,3}, \mathcal{R}_{2,3}): H \rightarrow 2S_5 + S_6, S_5 + S_6 \rightarrow P_1, S_5 \rightarrow P_2.$$

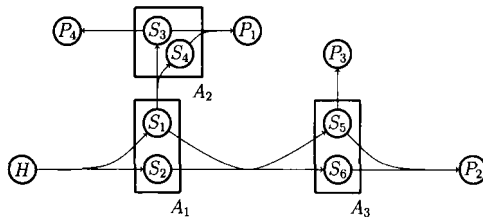


图 5 例 5.2 的强连通分解图例

由于三个子网络结构相同，只分析子网络 $(\mathcal{M}_{2,1}, \mathcal{R}_{2,1})$ 即可，其状态方程为

$$\begin{cases} \frac{d[S_1]}{dt} = 2v_1 - v_2 - v_3, \\ \frac{d[S_2]}{dt} = v_1 - v_2, \end{cases}$$

其中 $v_1 = u, v_2 = \frac{V_1[S_1][S_2]}{(k_1 + [S_1])(k_2 + [S_2])}, v_3 = \frac{V_2[S_1]}{k_3 + [S_1]}$.

为了使上面方程右端为零，需要 $v_1 = v_2 = v_3$ ，即

$$\begin{cases} \frac{V_2[S_1]}{k_3 + [S_1]} = u, \\ \frac{V_1[S_1][S_2]}{(k_1 + [S_1])(k_2 + [S_2])} = u. \end{cases} \quad (4)$$

由式 (4) 可解得

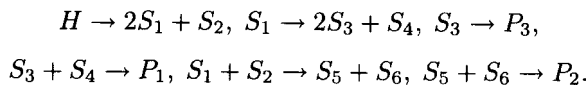
$$\begin{aligned} [S_1] &= \frac{k_3 u}{V_2 - u}, \\ [S_2] &= \frac{V_1 k_2 k_3}{k_3(V_1 - u) - k_1(V_2 - u)} - k_2. \end{aligned} \quad (5)$$

故这个子网络最多只有一组平衡态，即在这种拓扑结构下的网络不可能存在多个平衡态。由定理 4.4 知整体网络也不能产生多个平衡态。

对任意的输入 u ，可以调节参数 k_1, k_2, k_3, V_1, V_2 使 (4) 存在正解 $[S_1], [S_2]$ ，故 $(\mathcal{M}_{2,1}, \mathcal{R}_{2,1})$ 对于任意输入可维持平衡态。由定理 4.7 知整体网络对于任意输入可维持平衡态。

注 6 强连通分解把一个含有 6 个状态节点的整体网络分解成为 3 个含有 2 个节点的子网络。由于这些子网络结构类似，所以只需分析一个含有 2 个节点的子网络就可得到整体网络的性质。

例 5.3 考虑代谢网络



由强连通分解可以得到如下 3 个子网络

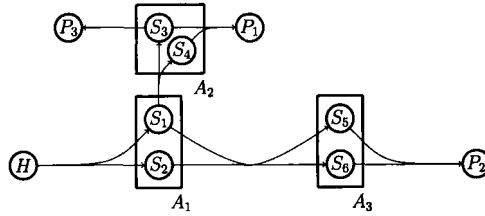
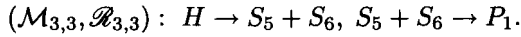
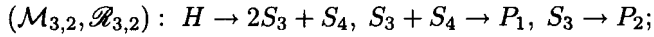
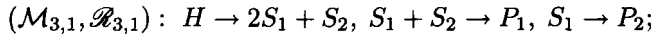


图 6 例 5.3 的强连通分解图例

对于子网络 $(\mathcal{M}_{3,3}, \mathcal{R}_{3,3})$, 其对应的状态方程为

$$\begin{cases} \frac{d[S_5]}{dt} = v_1 - v_2, \\ \frac{d[S_6]}{dt} = v_1 - v_2, \end{cases}$$

其中 $v_1 = u$, $v_2 = \frac{V_1[S_5][S_6]}{(k_1 + [S_5])(k_2 + [S_6])}$.

为了使上面方程的右端为零, 需要 $v_1 = v_2$, 即

$$u = \frac{V_1[S_5][S_6]}{(k_1 + [S_5])(k_2 + [S_6])}.$$

此时对应每一个 $u > 0$, 都存在参数 V_1, k_1, k_2 使得存在多个 $[S_5] > 0, [S_6] > 0$ 使得等式成立, 此时系统处于平衡状态. 所以该网络对于任意输入能产生多平衡态.

$(\mathcal{M}_{3,1}, \mathcal{R}_{3,1}), (\mathcal{M}_{3,2}, \mathcal{R}_{3,2})$ 与 $(\mathcal{M}_{2,1}, \mathcal{R}_{2,1})$ 结构相同, 由例 5.2 知其对于任意输入能维持多平衡态, 故由定理 4.6 可知, 整体网络对于任意输入能产生多个平衡态.

6 小 结

本文提出了一种强连通分解方法, 通过研究相对简单的子网络分析整体网络的多平衡态性质. 基于代谢网络的拓扑结构, 我们构造了其对应的相互作用图, 引入了紧缩运算的定义, 从而提出了强连通分解算法, 给出了该算法的时间复杂度, 证明了算法的合理性和可行性, 并阐明了子网络与整体网络在多平衡态性质意义下的关系.

本文仅是一个利用拓扑网络分析代谢网络平衡态的初步工作, 将来仍有许多有意义的问题需要考虑. 比如, 强连通子网络的定量分析, 结构更为一般的代谢网络的强连通分解等.

参 考 文 献

- [1] Kitano H. Systems biology: A brief overview. *Science*, 2002, **295**: 1662–1664.
- [2] Lei H B, Wang X, Wang R Q, Zhu X G, Chen L N, and Zhang J F. Analysis on steady states of photosynthetic carbon metabolic system. Proceedings of the Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, China, 2009.
- [3] Lei H B, Zhang J F, and Chen L N. Multi-equilibrium property of metabolic networks: Exclusion of multi-stability for SSN metabolic module. *International Journal of Robust and Nonlinear Control*, 2011, DOI:10.1002 /rnc.1718.
- [4] Su W, Lü J H, and Ogorzalek M J. Stability analysis of SSN biochemical network. Proceedings of the International Symposium on Circuits and Systems, Brazil, 2011.
- [5] Lei H B, Zhang J F, and Chen L N. Multi-equilibrium property of metabolic networks: SSI modul. *BMC Systems Biology*, <http://www.biomedcentral.com/1752-0509/5/S1/S15>, 2011.
- [6] Chen L, Wang R, and Zhang X. Biomolecular Network: Methods and Applications in Systems Biology. John Wiley & Sons, 2009.
- [7] Soule C. Graphic requirements for multistationarity. *ComplexUs*, 2003, **1**: 123–133.
- [8] Long S P, Zhu X G, Naidu S L, and Ort D R. Can improvement in photosynthesis increase crop yields. *Plant, Cell and Environment*, 2006, **29**: 315–330.
- [9] Pettersson G and Ryde-Pettersson U. A mathematical model of the calvin photosynthesis cycle. *Eur. J. Biochem.*, 1988, **175**: 661–672.
- [10] Poolman M G, Ölcer H, Lloyd J C, Raines C A, and Fell D A. Computer modelling and experimental evidence for two steady states in the photosynthetic calvin cycle. *Eur. J. Biochem.*, 2001, **268**: 2810–2816.
- [11] Zhu X G, Albab R, and de Sturler E. A simple model of the calvin cycle has only one physiologically feasible steady state under the same external conditions. *Nonlinear Analysis Series B: Real World Applications*, 2009, **10**: 1490–1499.
- [12] Stelling J, Klamt S, Bettenbrock K, Schuster S, and Gilles E D. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 2002, **420**: 190–193.
- [13] Schuster S, Pfeiffer T, Moldenhauer F, Koch I, and Dandekar T. Exploring the pathway structure of metabolism: Decomposition into subnetworks and application to *Mycoplasma pneumoniae*. *Bioinformatics*, 2002, **18**(2): 351–361.
- [14] Cinquin O and Demongeot J. Positive and negative feedback: Striking a balance between necessary antagonist. *J. Theor. Biol.*, 2002, **216**: 229–241.
- [15] Craciun G and Feinberg M. Multiple equilibria in complex chemical reaction networks: I. The injectivity property. *SIAM J. Appl. Math.*, 2005, **65**(5): 1526–1546.
- [16] Kaufman M, Soulé C, and Thomas R. A new necessary condition on interaction graphs for multi-stationarity. *J. Theor. Biol.*, 2007, **248**(4): 675–685.
- [17] Guo J and Zhang J F. Multi-equilibrium property of metabolic networks: MMN module. Proceedings of the 30th Chinese Control Conference, 2011.
- [18] Ma H W and Zeng A P. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics*, 2003, **19**(2): 270–277.
- [19] Schilling C H and Palsson B. Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *J. Theor. Biol.*, 2000, **203**: 249–283.
- [20] Ravasz E, Somera A L, Mongru D A, Oltvai Z N, and Barabasi A L. Hierarchical organization of modularity in metabolic networks. *Science*, 2002, **297**: 1551–1555.
- [21] 王冰, 修志龙, 唐焕文. 基于复杂网络理论的代谢网络结构研究进展. 中国生物工程杂志, 2006, **25**(6): 10–14.

- [22] Jensen J B and Gutin G. *Digraphs*. London: Springer, 2002.
- [23] Banaji M and Craciun G. Graphic-theoretic criteria for injectivity and unique equilibria in general chemical reaction systems. *Advances in Applied Mathematics*, 2010, **44**: 168–184.
- [24] Craciun G and Feinberg. Multiple equilibria in complex chemical reaction networks: I. The injectivity property. *SIAM J. Appl. Math.*, 2005, **65**(5): 1526–1546.
- [25] Michaelis L and Menten M. Die kinetic der invertinwirkung. *Biochem. Z.*, 1913, **49**: 333–369.
- [26] Monod J, Wyman J, and Changeux J P. On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology*, 1965, **12**: 88–118.
- [27] Tarjan R. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1972, **1**(2): 146–160.

A STRONG CONNECTIVITY DECOMPOSITION METHOD FOR ANALYZING THE MULTI-EQUILIBRIUM PROPERTY OF GENERAL METABOLIC NETWORKS

BI Wenjian GUO Jin ZHAO Yanlong ZHANG Jifeng

(*Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190*)

Abstract For general metabolic networks whose reaction rates satisfy some conditions, a strong connectivity decomposition (SCD) method is proposed, which can not only divide the whole network into a set of sub-networks but also keep the strong connectivity of the network. This makes it possible to understand the multi-equilibrium property of the whole network by analyzing the sub-networks. The SCD method is based on only the topological structure of the network. To get an SCD for a given metabolic network, the concepts of metabolic reaction graph, interaction graph and contraction operation are introduced. It is shown that for a given metabolic network, the SCD is unique, all the sub-networks are strongly connected, and the computational complexity of the decomposition is polynomial. The relationship between the whole network and sub-networks is given in the sense of multi-equilibrium properties. Examples are given to demonstrate the effectiveness of the algorithms and the main results.

Key words Metabolic network, strong connectivity decomposition, interaction graph, Multi-Equilibrium property, contraction operation.